

SISTEM ZA PREPOZNAVANJE ŠTAMPANOG TEKSTA KORIŠĆENJEM TESSERACT BIBLIOTEKE**SYSTEM FOR PRINTED TEXT RECOGNITION USING TESSERACT LIBRARY**Mikloš Popović, *Fakultet tehničkih nauka, Novi Sad***Oblast – MEHATRONIKA**

Kratak sadržaj – U radu je predstavljeno jedno rešenje sistema koji omogućava prepoznavanje štampanog teksta na slici upotrebom digitalne obrade slike, tj. mašinske vizije. Cilj je bio napraviti aplikaciju pisanjem programskog koda u Python-u uz upotrebu Tesseract biblioteke koja omogućava prepoznavanje štampanog teksta, nakon čega se formira fajl u kome se nalazi prepoznati tekst. Analizirana je osetljivost predloženog rešenja na način akvizicije pomoću kamere, tj. veličinu slova, kao i koje tehnike obrade najbolje utiču na željeni rezultat.

Ključne reči: Prepoznavanje štampanog teksta, Tesseract, Python, mašinska vizija

Abstract – Paper presents one solution of the system that allows recognition of the printed text on the image using digital image processing, i.e. machine vision. The goal was to create an application by writing the program code in Python and by using the Tesseract library, which allows printed text recognition, after which a file containing the recognized text is formed. Sensitivity of the proposed solution was analyzed against the type of acquisition using camera, i.e. character size, as well as which processing techniques best affect the desired result.

Keywords: Printed text recognition, Tesseract, Python, machine vision

1. UVOD

Osnovni problemi efikasnog prepoznavanja znakova su složena heterogena pozadina, promenljivost pisanja znakova, niska rezolucija slike, razlike u kontrastu i osvetljenosti, prisustvo šuma, izobličenja i slično.

Tekst je skup simbola koji se koriste za, komunikaciju ili nasleđivanje kulture. Kao jedan od izuma čovečanstva koji ima najveći uticaj, tekst je odigrao važnu ulogu u životu ljudi. Mašinska vizija je tehnologija i skup metoda koje se koriste za pružanje automatskog pregleda i analize utemeljene na slikama i nalazi primenu u automatskom pregledu, kontroli procesa i vođenja robota u industriji. Viziju kao alat koriste mnoge tehnologije, integrisani sistemi softverski i hardverski proizvodi.

NAPOMENA:

Ovaj rad proistekao je iz master rada čiji mentor je bio dr Branko Brkljač, doc.

2. MAŠINSKA VIZIJA I PRAKTIČNI PRIMERI

Mašinska vizija je upotreba uređaja za optičko beskontaktno detektovanje objekata gde se automatski prima i tumači slika dobijena iz realne scene, a u cilju dobijanja željene informacije ili upravljanja mašinom i nekim procesom. Sistem mašinske vizije koristi kameru za analizu slike i uglavnom je ograničen na određeni specifičan zadatak za koji se očekuje da će se obavljati u relativno kontrolisanim uslovima, dok računarska ili kompjuterska vizija razmatra i širu klasu problema, odnosno algoritme i postupke namenjene obradi i interpretaciji slike u nekontrolisanim okruženjima.

Čovek vidi uz pomoć elektromagnetskog zračenja u opsegu vidljive svetlosti. Svetlost pada na objekte, reflektuje se i mi ih „vidimo“. Tako, vidljiva svetlost omogućava izradu različitih slika. Međutim, slika ne nastaje samo pomoću vidljive svetlosti.

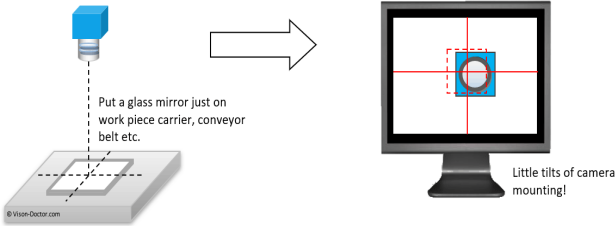
Druge vrste zračenja (gama, rendgensko, ultralubičasto, infracrveno, akustično), takođe mogu da proizvedu sliku primenom odgovarajućih senzora. Ono što dodatno karakteriše ljudski vid jeste da se ovo čulo lako prilagođava na promene, čovek je sposoban da vrlo brzo razume i da se snađe u kompleksnim scenama, ali je ljudsko oko ograničeno na vidljiv spektar svetlosti.

Nasuprot tome, mašinsku viziju karakteriše konzistentnost, ne postoji zamor koji je prisutan kod čoveka, može da funkcioniše čak i u infracrvenom spektru i sa X zracima. Najbitnija prednost nad ljudskim vidom jeste da mašinska vizija strogo i precizno prati unapred zadati algoritam po kojem funkcioniše i lako se prilagođava 'neprijateljskim' sredinama u kojima čovek ne bi mogao da opstane.

Mašinska vizija se na primer koristi za:

- tehnologiju brze proizvodne linije
- mikroskopski pregled,
- rad u opasnom okruženju
- kontroli procesa u zatvorenom krugu
- navođenje robota,
- precizno bezkontaktno merenje
- poravnanje položaja kamere (Slika 1)

The „trick with the mirror“:



Slika 1. Primer poravnanja kamere [1]

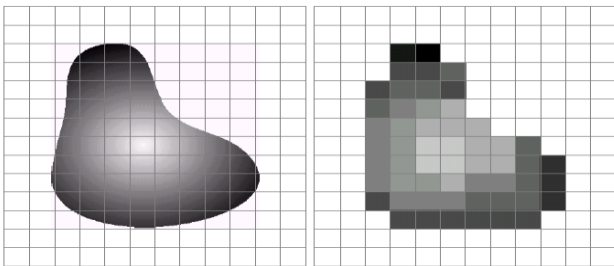
3. POJAM SLIKE

Digitalna slika se može definisati kao skalarna ili vektorska funkcija dve prostorne promenljive koja je generisana optičkim sredstvima, uzorkovana u jednako razmaknutim tačkama i kvantizovana jednakim intervalima amplitude. Prema tome, digitalna slika se predstavlja 2D matricom kvantizovanih vrednosti.

Prve slike dobijene fotografskim postupkom (hemikalije osetljive na svetlost, nanete na neku podlogu) datiraju oko 1838. godine.

Prve digitalne slike sa potrošačkom elektronikom nastaju oko 1990. godine, pojavom Kodak digitalne kamere. Danas digitalna slika nastaje ili digitalizacijom klasičnih slika, ili direktno, pomoću digitalne kamere.

Digitalizacijom analogne električne veličine dobija se digitalna slika. Generalno postoje tri tipa senzora: tačkasti, linijski i matični. Na slici 2. dat je primer digitalizacije matičnim senzorom.



Slika 2. Digitalizacija slike [2]

3.1. Obrade i operacije nad slikom

Dobijena slika se često ne može upotrebiti, pa se javlja potreba za poboljšanjem slika, izoštravanjem, isticanjem kontrasta i slično. U nastavku će biti pomenute neke od bitnih operacija i načina obrada slike [2, 3].

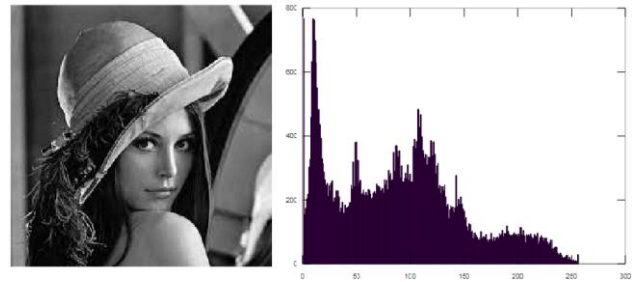
a) Osnovne aritmetičke operacije koje se koriste u poboljšanju slike:

- o Oduzimanje
- o Sabiranje
- o Množenje
- o Deljenje

b) Izračunavanje histograma

Da bi se što lakše donela odluka o tome kakva je transformacija ulaznih u izlaznu sliku potrebna (da bi se definisalo preslikavanje koje omogućava poboljšanje), često se koristi histogram slike, odnosno tehnika ekvalizacije histograma.

Zastupljenost pojedinih piksela u slici lako je uočljiva na osnovu histograma, Slika 3.



Slika 3. Slika Lena i histogram

c) Ublažavanje slike

Ostvaruje se usrednjavanjem (integracijom) slike. Efekat usrednjavanja može da bude redukcija šuma. Šum predstavlja nagle (oštre) promene osvetljaja u slici.

Međutim, ivice u slici takođe predstavljaju promene osvetljaja, pa se ublažavanjem slike gubi na nivou detalja, što je jedan od neželjenih efekata ublažavanja. Ublažavanje može da ima za cilj i zamućivanje slike, tj. preprocesiranje slike kojim se ukidaju sitni detalji pre ekstrakcije velikih objekata.

d) Izoštravanje (engl. sharpening) je obrnut proces od ublažavanja. Cilj izoštravanja je naglašavanje finih detalja u slici i otklanjanje zamućenja koje je nastalo usled greške ili zbog prirode sistema za akviziciju slike.

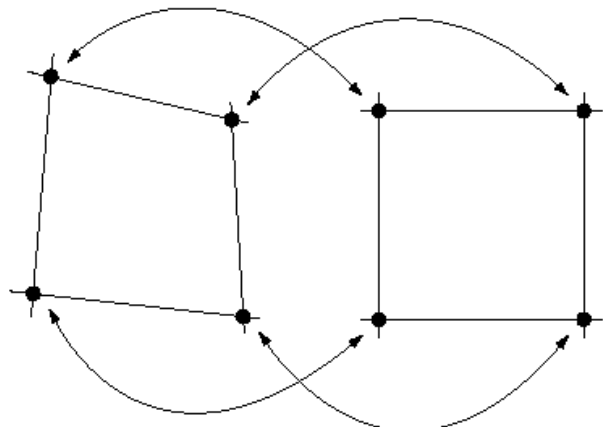
Pošto se ublažavanje slike ostvaruje usrednjavanjem (integracijom), izoštravanje se realizuje prostornim diferenciranjem slike. Jedan primer je dat na Slici 4.



Slika 4. Rezultat nakon izoštravanja

Prostorna transformacija je transformacija koja preslikava lokacije piksela iz ravni originalne slike u ravan degradirane (izobličene) slike. Najčešće se prostorna relokacija piksela u okviru jednog regiona definiše na osnovu povezanih tačaka i relacija između njih. Povezane tačke su granice regiona.

Primer prostorne transformacije prikazan je na Slici 5. Svaka označena tačka sa levog, izvitoperenog kvadrata, dobija mesto na desnom, a rezultat je pravilan kvadrat.



Slika 5. Prostorna transformacija

4. PYTHON I EKSPERIMENTI

Za ovaj rad odabran jezik je Python. Python je interpreterski, objektno-orijentisani jezik visokog nivoa, namenjen za pravljenje svih vrsta aplikacija – od inženjerskih i naučnih, do poslovnih i web primena. Jezik je osmislio holandski programer Guido van Rossum krajem osamdesetih godina 20. veka.

Pycharm je namensko Python integrisano razvojno okruženje koje pruža širok spektar osnovnih alata za Python programere, tesno integrisanih da bi se stvorilo pogodno okruženje za produktivan rad.

4.1 OpenCV

OpenCV (engl. *Open Source Computer Vision*) je biblioteka otvorenog koda koja je nastala kao deo alata u istraživačkoj laboratoriji kompanije Intel. Biblioteka je u međuvremenu doživela značajna unapređenja i razvoj, da bi se danas koristila kao čest izbor na različitim platformama i kao alat u mnogim aplikacijama na različitim operativnim sistemima.

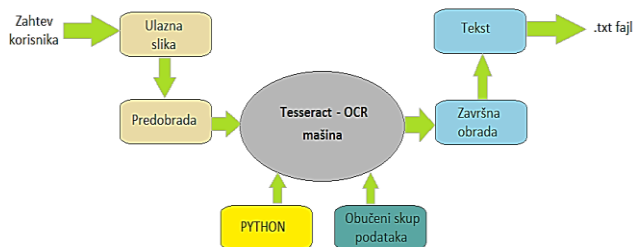
Osnovna prednost OpenCV-a je pružanje jednostavnog okruženja za rad sa mašinskom vizijom i omogućavanje jednostavnog i brzog razvoja algoritama, a time i aplikacija koje koriste mašinsku viziju.

4.2 Tesseract

Najveći izazov tokom razvoja predloženog rešenja je bio kako da se od slike koja sadrži štampani tekst, prepozna pa izdvoji samo tekst, tj. kako od digitalne slike da se dobije tekst koji se u nekim kasnijim koracima može koristiti na različite načine u domenu robotike i automatizacije. Prepoznavanje i izdvajanje teksta sa digitalne slike omogućeno je Tesseract bibliotekom, koja će ovde biti detaljnije objašnjena.

Tesseract je „mašina“ za optičko prepoznavanje karaktera sa otvorenim kodom i predstavlja jednu od najpopularnijih i najkvalitetnijih OCR biblioteka koje koristi veštačku inteligenciju za pretragu teksta i njegovo prepoznavanje na slikama.

Tesseract pronalazi šablone u pikselima, slovima, rečima i rečenicama [4]. Koristi pristup u dva koraka koji se naziva adaptivno prepoznavanje. Prva faza predstavlja prepoznavanje karaktera, dok u drugoj fazi program ispunjava slova u koja nije siguran slovima koja mogu da odgovaraju kontekstu reči ili rečenice.



Slika 6. Proces prepoznavanja teksta

Proces prepoznavanja teksta, počev od ulazne slike, pa do kraja procesa, tj. dobijanja tekstualnog fajla sa prepoznatim tekstom, prikazan je na Slici 6.

4.3 Eksperimenti prepoznavanja teksta

U ovom delu biće prikazani rezultati koji su dobijeni nakon nekoliko eksperimenata koji su nazvani prema veličini slova kojim je tekst napisan. Eksperiment podrazumeva da je papir sa tekстом prošao kroz program obrade, izoštravanja i nakon takve dobijene slike treba da se dobije prepoznati štampani tekst. Prikazani su rezultati slika koje su dobijene fotografisanjem sa kamerom postavljenom na fiksni držač, dok je druga vrsta slika obuhvatila fotografije dobijene sa kamerom koja nije nepokretna, Slika 7.

Tekst koji treba da se prepozna je sledeći:

**MIKLOS
POPOVIC 72
Sistem za
prepoznavanje
stampanog
teskta**

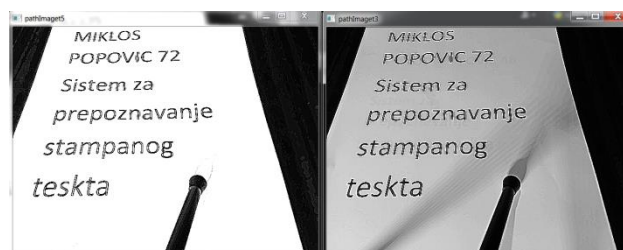


Slika 7. Eksperiment 'font 48'

Leva slika je jasnija jer je prošla kroz dodatnu obradu filterom za izoštravanje. Rezultat dobijenog prepoznatog teksta dat je u Tabeli 2., a tabelom će biti prikazani i ostali rezultati.

Tabela 2. Rezultat prepoznavanja teksta sa fontom 48

Slika 5	Slika 3
I) pathimaget5 a= eRe	pathimaget3
MIKLOS POPOVIC 48 Sistem za prepoznavanje stampanog teskta	MIKLOS POPROMIC 48 Sistem za prepoznavanje stampanog tesktay



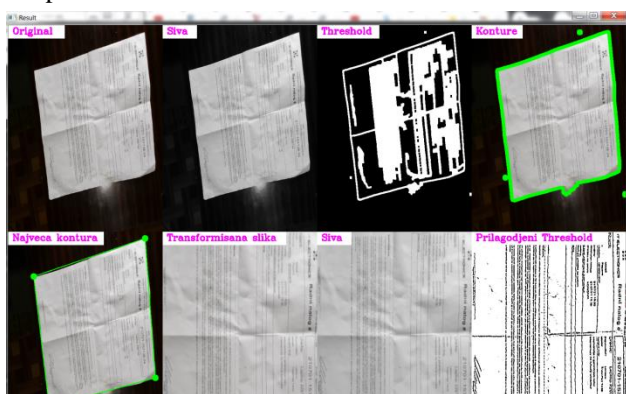
Slika 8. Eksperiment 'font 72'

Na Slici 8 veličina slova je veća u odnosu na prošli eksperiment, pa su i rezultati mnogo bolji nego prethodni, što se vidi iz tabele 3.

Tabela 3. Rezultat 'font 72'

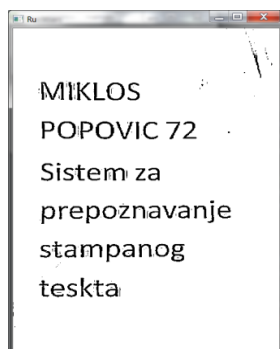
Slika 5	Slika 3
MIKLOS POPOVIC 72	MIKLOS POPOVIC 72
Sistem za prepoznavanje stampanog teskta	Sistem za prepoznavanje stampanog teskta sg

Sledeći eksperiment ima dodatnu obradu koja se zove prostorna transformacija. Papir se postavi na sto ili zalepi za zid. Kamera se prinese tako da ceo papir bude 'uhvaćen' na monitoru tako da program sam pronalazi najveću konturu, ocrta je i na 5. delu Slike 9 je izdvaja. Ovako dobijena slika je čitljivija nego slikanjem sa stalka bez pomenute obrade.



Slika 9. Fotografija dobijena nakon prostorne transformacije

Sledeća slika daje najbolje rezultate prepoznavanje teksta.



Slika 10. Primer fotografije dobijen nakon prostorne transformacije

Ovakva slika, kao što je ilustrovano na Slici 10 daje najbolje rezultate prepoznavanja teksta. Reči su čitljive nema izobličenja, otklonjene su senke, kontrast je jasan, a rezultat bolji, tj. dobijeni tekst nakon prepoznavanja identičan je onom tekstu koji se nalazi na papiru.

5. ZAKLJUČAK

U ovom radu demonstrirana je primena Tesseract biblioteke za prepoznavanje štampanog teksta. Dobijeno je zadovoljavajuće rešenje za traženi proces. Na osnovu testiranja predlaže se da se prilikom korišćenje ove biblioteke takođe primene i dodatne tehnike predobrade slike kao što su: ublažavanje, izoštravanje, prostorna transformacija i morfološke operacije.

Jedan od uočenih nedostataka jeste da sa primenjenom kamerom, na datom rastojanju, ne može da se prepozna tekst standardne veličine odnosno fonta 12. Za prepoznavanje teksta takve veličine predlaže se kamera boljeg kvaliteta ili neki napredniji algoritam prepoznavanja, odnosno predobrade.

U radu su pored izrade aplikacije opisani primeri mašinske vizije, dok je predloženo rešenje moguće koristiti kao deo različitih robotizovanih platformi.

6. LITERATURA

- [1] Chen, X., Jin, L., Zhu, Y., Luo, C., & Wang, T. (2021). Text recognition in the wild: A survey. *ACM Computing Surveys (CSUR)*, 54(2), 1-35.
- [2] Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing*, Hoboken.
- [3] Sonka, M., Hlavac, V., & Boyle, R. (2014). *Image processing, analysis, and machine vision*. Cengage Learning.
- [4] Smith, R. (2007) An overview of the Tesseract OCR engine. *ICDAR 2007*, Vol. 2, pp. 629-633. IEEE.

Kratka biografija:



Mikloš Popović rođen je u Apatinu 1986. god. Diplomski rad iz oblasti Programiranja logičkih kontrolera – Upravljanje pristupom prostora sa ulazno izlaznim barijerama, odbranio je 2018. godine na Fakultetu tehničkih nauka u Novom Sadu.